

AMENDMENT TO THE CLAIMS

~~1~~ (currently amended) A method of processing a body of text to generate compression options, comprising:

performing a linguistic analysis on the body of text to obtain a linguistic output indicative of linguistic components of the body of text; and

~~after performing the linguistic analysis, generating automatically generating a plurality of~~
compression options for each of a plurality of different portions of the body of text to compress the body of text based on the linguistic output each of the compression options comprising a different compressed form of an instance of the portion in the body of text; and

~~selecting one of the plurality of compression options for each of the plurality of different~~
~~portions of the body of text to output a compressed form of the body of text.~~

~~2~~ (previously amended) The method of claim 1 wherein generating a plurality of compression options comprises:

automatically subjecting the portions of the body of text to different sets of compression rules to obtain the plurality of compression options.

~~3~~ (previously amended) The method of claim 2 wherein automatically subjecting the portions of the body of text to different sets of compression rules, comprises:

subjecting each portion of the body of text to the different sets of compression rules in a predetermined order such that the compression options reflect varying degrees of compression of a same portion of the body of text.

~~4~~ (previously amended) The method of claim 3 wherein generating automatically a plurality of compression options comprises:

generating a compression identifier attribute indicative of at least one of the sets of compression rules to which the portion of the body of text is subjected.

09/18/823

~~5.~~ (previously amended) The method of claim 4 wherein generating automatically a plurality of compression options comprises:

generating a ShortForm attribute indicative of a compressed form of the portion of the body of text after application of the set of compression rules.

~~6.~~ (previously amended) The method of claim 5 wherein generating automatically a plurality of compression options comprises:

generating a case normalized attribute, based on the ShortForm attribute, indicative of a CaseNormalizedForm of the ShortForm attribute.

~~7.~~ (previously amended) The method of claim 6 wherein generating automatically a plurality of compression options comprises:

generating a compression attribute indicative of a further compressed form of the case normalized attribute.

~~8.~~ (original) The method of claim 7 wherein generating a compression attribute comprises:

applying letter removal rules to the case normalized attribute to remove letters based on a predetermined location of the letters in the CaseNormalizedForm.

~~9.~~ (previously amended) The method of claim 8 wherein generating automatically a plurality of compression options comprises:

generating a LongForm attribute that reflects substantially no compression of the portion of the body of text.

~~10.~~ (previously amended) The method of claim 9 wherein one ShortForm attribute comprises a word substitution based on a dictionary look-up and wherein generating automatically a plurality of compression options comprises:

setting the case normalized attribute and the compression attribute to the ShortForm

attribute.

~~11.~~ (original) The method of claim 5 wherein performing a linguistic analysis comprises performing a syntactic analysis on the portion of the body of text and wherein generating the ShortForm attribute comprises:

applying the set of compression rules based on the syntactic analysis.

~~12.~~ (original) The method of claim 11 wherein the linguistic analysis further comprises, prior to performing the syntactic analysis:

performing a lexical analysis on the body of text; and

performing a morphological analysis on the body of text.

~~13.~~ (original) The method of claim 5 wherein generating the ShortForm attribute comprises:
normalizing dates to a numerical form.

~~14.~~ (original) The method of claim 5 wherein generating the ShortForm attribute comprises:
normalizing offset dates to a numerical form, based on a date that the body of text was authored.

6,026,410 Allen et al
Col. 16, 24 - 17, 60

~~15.~~ (original) The method of claim 5 wherein generating the ShortForm attribute comprises:
maintaining symbol-sensitive text fragments in uncompressed form.

to provide a
means of ~~access~~
interpreting data
from national
language
(Col. 16
24-27)

~~16.~~ (original) The method of claim 15 wherein maintaining symbol-sensitive text fragments comprises:

maintaining text fragments that, cannot be accurately understood unless maintained fully in-tact, in uncompressed form.

~~17.~~ (original) The method of claim 16 wherein maintaining text fragments comprises:

maintaining uniform resource locators and electronic mail addresses in uncompressed form.

~~18.~~ (original) The method of claim 11 wherein the syntactic analysis includes a tree having non-terminal nodes representing multi-word portions of the body of text and terminal nodes indicative of words in the body of text, and wherein both the non-terminal nodes and the terminal nodes are examined for application of compression rules.

~~19.~~ (currently amended) A computer readable data structure formed from a linguistic analysis of a body of text to be compressed indicative of a plurality of compressed forms of the body of text, the data structure comprising:

a plurality of different sections, each section corresponding to a textual term in the body of text, each section further comprising a plurality of selectable data fields, selectable to representing a plurality of different compressed forms of the corresponding textual term in the body of text.

~~20.~~ (previously amended) The computer readable data structure of claim 19 and further comprising:

a compression type attribute indicative of a type of compression applied to the textual term in the body of text in generating at least one of the plurality of compressed forms.

~~21.~~ (previously amended) The computer readable data structure of claim 20 wherein the plurality of compressed forms comprises:

a ShortForm attribute indicative of a compressed form of the textual term in the body of text after application of the type of compression identified by the compression type attribute.

~~22.~~ (previously amended) The computer readable data structure of claim 21 wherein the plurality of compressed forms comprises:

a case normalized attribute, based on the ShortForm attribute, indicative of a CaseNormalizedForm of the ShortForm attribute.

~~23.~~ (previously amended) The computer readable data structure of claim 22 wherein the plurality of compressed forms comprises:

a compression attribute indicative of a further compressed form of the case normalized attribute.

~~24.~~ (previously amended) The computer readable data structure of claim 23 and further comprising:

a LongForm attribute indicative of substantially no compression of the textual term in the body of text.

~~25.~~ (currently amended) A message handler receiving a message and generating compression options indicative of different forms of a portion of a body of text in the message, the message handler comprising:

a linguistic analyzer linguistically configured to analyze the body of text and provide a linguistic analysis; and

a compression form generator configured to automatically generate a plurality of different compressed forms of a plurality of individual textual segments in the body of text based on the linguistic analysis; and

a compressor configured to generate an output indicative of selected ones of the plurality of different compressed forms for the individual textual segments in the body of text.

~~26.~~ (previously amended) The message handler of claim 25 wherein the compression form

generator is configured to automatically apply a plurality of different sets of compression rules to each of the individual textual segments in the body of text to obtain the plurality of compressed forms.

~~27.~~ (previously amended) The message handler of claim 26 wherein the compression form generator is further configured to automatically apply the different sets of compression rules in a predetermined order such that the plurality of compressed forms reflect varying degrees of compression of a same individual textual segment in the body of text.

~~28.~~ (previously amended) The message handler of claim 27 wherein the compression form generator is further configured to generate a compression identifier attribute indicative of at least one of the sets of compression rules applied to the individual textual segment in the body of text.

~~29.~~ (original) The message handler of claim 27 wherein the compression form generator is configured to provide, at its output, a data structure containing a plurality of attributes indicative of the plurality of compressed forms, and the compression identifier attribute.

~~30.~~ (previously amended) The message handler of claim 29 wherein the plurality of attributes includes:

- a ShortForm attribute indicative of a compressed form of the individual textual segment in the body of text after application of the set of compression rules;
- a case normalized attribute, based on the ShortForm attribute, indicative of a CaseNormalizedForm of the ShortForm attribute; and
- a compression attribute indicative of a further compressed form of the case normalized attribute.

~~31.~~ (previously amended) The message handler of claim 30 wherein the plurality of attributes further comprises:

a LongForm attribute that reflects substantially no compression of the individual textual segment in the body of text.